(54) Title: POLARISATION MODE DISPERSION COMPENSATOR USING A MEAN SQUARE ERROR TECHNIQUE
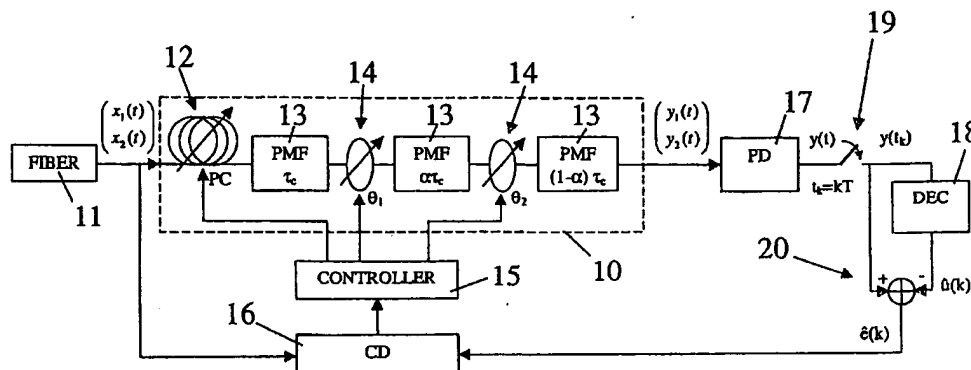
(57) Abstract: A method for the adaptive adjustment of a PMD compensator in optical fiber communication systems comprises the steps of taking the signal at the compensator output and extracting the components $y_1(t)$ and $y_2(t)$ on the two orthogonal polarizations, computing the signal $y(t)=[y_1(t)]^2+[y_2(t)]^2$, sampling the signal $y(t)$ at instants $t_k=kT$ with T = symbol interval to obtain samples $y(t_k)$, computing the mean square error $e(k)=y(t_k)-u(k)$ with $u(k)$ equal to the symbol transmitted, and adjusting the parameters of the compensator to seek to minimize $e(k)$. A PMD compensator in accordance with the method comprises an adjustment system which comprises in turn a photodetecor (17) which takes the components $y_1(t)$ and $y_2(t)$ on the two orthogonal polarizations from the signal at the compensator output, a sampler (19) which samples at instants $t_k=kT$ with T = symbol interval, the signal $y(t)=[y_1(t)]^2+[y_2(t)]^2$ at the output of the photodetector (17) to obtain samples $y(t_k)$, a circuit (18, 20) for computation of the mean square error $e(k)=y(t_k)-u(k)$ with $u(k)$ equal to the symbol transmitted, and a regulator (15, 16) which regulates the parameters of the compensator to seek to minimize $e(k)$.

POLARISATION MODE DISPERSION COMPENSATOR USING A MEAN SQUARE ERROR TECHNIQUE

The present invention relates to methods of adaptive adjustment of PMD compensators in optical fiber communication systems. The present invention also relates to a compensator in accordance with said method.

5   In optical fiber telecommunications equipment the need to compensate the effects of polarization mode dispersion (PMD) which occur when an optical signal travels in an optical fiber based connection is known.

It is known that PMD causes distortion and dispersion of optical signals sent over

10   optical fiber connections making the signals distorted and dispersed. The different time delays among the various signal components in the various polarization states acquire increasing importance with the increase in transmission speeds. In modern optical fiber based transmission systems with ever higher frequencies (10 Gbit/s and more), accurate compensation of PMD effects becomes very important and delicate. This compensation

15   must be dynamic and performed at adequate speed.

The general purpose of the present invention is to remedy the above mentioned shortcomings by making available a method of fast, accurate adaptive adjustment of a PMD compensator and a compensator in accordance with said method.

In view of this purpose it was sought to provide in accordance with the present invention a method for the adaptive adjustment of a PMD compensator in optical fiber communication systems with the compensator comprising a cascade of adjustable optical devices over which passes an optical signal to be compensated comprising the steps of extracting the y1(t) and y2(t) on the two orthogonal polarizations of the signal at the compensator output, obtaining the signal $y(t)=|y_1(t)|^2+|y_2(t)|^2$, sampling the signal $y(t)$ at instants $t_k=kT$ with T = symbol interval to obtain samples $y(t_k)$, computing the mean square error $e(k)=y(t_k)-u(k)$ with $u(k)$ equal to the symbol transmitted or with $u(k)$ replaced by a decision $\hat{u}(k)$ on the transmitted symbol $u(k)$, and producing control signals for parameters of at least some of said adjustable optical devices to tend toward minimization of $e(k)$.

In accordance with the method it was also sought to realize a PMD compensator in optical fiber communication systems applying the method and comprising a cascade of adjustable optical devices over which passes an optical signal to be compensated and an adjustment system comprising a photodetector (17) which takes the components $y_1(t)$ and $y_2(t)$ on the two orthogonal polarizations from the signal at the compensator output, a sampler 19 which samples at instants $t_k=kT$ with T = symbol interval, the signal $y(t)=|y_1(t)|^2+|y_2(t)|^2$ at the output of the photodetector 17 to obtain samples $y(t_k)$, a circuit 18,20 for computation of the mean square error $e(k)=y(t_k)-u(k)$ with $u(k)$ equal to the symbol transmitted and a regulator 15,16 which regulates parameters of at least some of said optical devices to tend towards minimization of $e(k)$.

To clarify the explanation of the innovative principles of the present invention and its advantages compared with the prior art there is described below with the aid of the annexed drawings a possible embodiment thereof by way of non-limiting example applying said principles. In the drawings –

5

The components $y_1(t)$ and $y_2(t)$ on the two orthogonal polarizations are sent

FIG 1 shows a block diagram of a PMD compensator with associated control circuit, and

10

FIG 2 shows an equivalent model of the PMD compensator.

With reference to the FIGS FIG 1 shows the structure of a PMD compensator designated as a whole by reference number 10. This structure consists of the cascade of some optical devices which receive the signal from the transmission fiber 11. The first

15 optical device is a polarization controller 12 (PC) which allows modification of the optical signal polarization at its input. Thus there are three polarization maintaining fibers 13 (PMF) separated by two optical rotators 14.

20 A PMF fiber is a fiber which introduces a predetermined differential group delay (DGD) between the components of the optical signal on its two principal states of polarization (PSP) termed slow PSP and fast PSP.

In the case of the compensator shown in FIG 1 the DGD delays at the frequency of the optical carrier introduced by the three PMFs are respectively τ_c, $\alpha\tau_c$ and $(1-\alpha)\,\tau_c$, with $0<\alpha<1$ and with τ_c and α which are design parameters.

5   An optical rotator is a device which can change the polarization of the optical signal upon its input by an angle θ_i (the figure shows θ_1 for the first rotator and θ_2 for the second) on a maximum circle on the Poincarè sphere.

An optical rotator is implemented in practice by means of a properly controlled PC.

10

In FIG 1, $x_1(t)$ and $x_2(t)$ designate the components on the two PSPs of the optical signal at the compensator input whereas similarly $y_1(t)$ and $y_2(t)$ are the components of the optical signal at the compensator output.

15   The components $y_1(t)$ and $y_2(t)$ on the two orthogonal polarisations are sent to the input of the photo detector (PD) which produces a signal $y(t)$ given by:

$$y(t)=|y_1(t)|^2+|y_2(t)|^2 \tag{1}$$

20   This signal can if necessary be filtered by a post-detection filter. Without loss of generality let us assume that this filter is not present. If present this filter produces obvious changes in the adaptive adjustment of the compensator parameters. These changes are readily imaginable to those skilled in the art and are therefore not further described.

As shown again in FIG 1 the signal y(t) is then sampled by a sampler 19 at the instants $t_k=kT$, where T is the symbol interval. Based on the sample $y(t_k)$ a decision on the transmitted bit u(k) is made. Let us designate here by û(k) this decision made by a known decider circuit 18 (DEC).

5

The input-output behavior of each of the optical devices is described here by means of the so-called Jones transfer matrix H(w) which is a 2x2 matrix characterized by frequency-dependent components. Denoting by $W_1(\omega)$ and $W_2(\omega)$ the Fourier transforms of the components of the optical signal at the device input the Fourier

10   transforms $Z_1(\omega)$ and $Z_2(\omega)$ of the components of the optical signal at the device output are given by:

$$\begin{pmatrix} Z_1(\omega) \\ Z_2(\omega) \end{pmatrix} = H(\omega)\begin{pmatrix} W_1(\omega) \\ W_2(\omega) \end{pmatrix} \tag{2}$$

15   Thus the Jones transfer matrix of the PC is

$$\begin{pmatrix} h_1 & h_2 \\ -h_2^* & h_1^* \end{pmatrix} \tag{3}$$

where h_1 e h_2 satisfy the condition $|h_1|^2+|h_2|^2 =1$ and are frequency independent.

20

Denoting by ϕ_1 and ϕ_2 the PC control angles, h_1 and h_2 are expressed by:

$$h_1=-\cos(\phi_2-\phi_1)+ j \sin(\phi_2-\phi_1) \sin\phi_1 \tag{4}$$

$$h_2=-j \sin(\phi_2-\phi_1) \cos\phi_1$$

Clearly if the PC is controlled using other angles or voltages, different relationships will

correlate these other parameters with h_1 and h_2. The straightforward changes in the

algorithms for adaptive adjustment of the PMD compensator are discussed below.

5   Similarly, an optical rotator with rotation angle θ_i is characterized by the following

Jones matrix:

$$\begin{pmatrix} \cos\theta_i & \sin\theta_i \\ -\sin\theta_i & \cos\theta_i \end{pmatrix} \tag{5}$$

The Jones transfer matrix of a PMF with DGD τ_i may be expressed as RDR^{-1} where D is

10   defined as:

$$D \triangleq \begin{pmatrix} e^{j\omega\tau_i/2} & 0 \\ 0 & e^{-j\omega\tau_i/2} \end{pmatrix} \tag{6}$$

and R is a unitary rotation matrix accounting for the PSPs' orientation. This matrix R

15   may be taken as the identity matrix I without loss of generality when the PSPs of all the

PMFs are aligned.

As shown in FIG 1, to control the PMD compensator a controller 15 is needed to

produce optical device control signals of the compensator calculated on the basis of the

20   quantities sent to it by a controller pilot 16 termed controller driver (CD).

The CD feeds the controller with the quantities needed to update the compensator

optical device control parameters. As described below, these quantities will be

extracted by the CD from the signals at the input and/or output of the compensator.

The controller will operate following the criterion described below and will use one of the two algorithms described below.

To illustrate the PMD compensator adaptive adjustment algorithms let us assume that the controller can directly control the parameters ϕ_1, ϕ_2, θ_1 and θ_2 which we consolidate in a vector θ defined as:

$$\theta \triangleq (\phi_1, \phi_2, \theta_1, \theta_2)^T$$

If it is not so, in general there will be other parameters to control, for example some voltages, which will be linked to the previous ones in known relationships. The simple changes necessary are discussed below.

The PMD being a slowly changing phenomenon adjustment of the compensator parameters will be performed at a rate lower then the transmitted symbol rate 1/T. Let us assume that this adjustment is performed at the discrete-time instants $t_nL=nLT$ where $L\geq1$. We designate by:

$$\theta(nL)=(\phi_1(nL), \phi_2(nL), \theta_1(nL), \theta_2(nL))^T$$

the value of the compensator parameters after the nth update.

In accordance with the method of the present invention the compensator parameter adjusting criterion uses the mean square error (MSE) criterion.

Based thereon, the compensator parameters θ are adjusted to minimize the mean square error e(k) defined as:

$$e(k) = y(t_k) - u(k) \qquad (7)$$

5

This error is a function of θ through $y(t_k)$. We explicit this dependence by defining $F(\theta) = e(k)$. Therefore, the performance index to be minimized is the mean value of $F^2(\theta)$. The compensator parameters θ will be updated by the rule:

10

$$\phi_1[(n+1)L] = \phi_1(nL) - \gamma \frac{\partial E\{F^2(\theta)\}}{\partial \phi_1}\bigg|_{\theta=\theta(nL)} = \phi_1(nL) - 2\gamma F(\theta)\frac{\partial E\{F(\theta)\}}{\partial \phi_1}\bigg|_{\theta=\theta(nL)} =$$

$$= \phi_1(nL) - 2\gamma[y(t_{nL}) - u(nL)]\frac{\partial E\{y(t_{nL})\}}{\partial \phi_1}$$

$$\phi_2[(n+1)L] = \phi_2(nL) - \gamma \frac{\partial E\{F^2(\theta)\}}{\partial \phi_2}\bigg|_{\theta=\theta(nL)} = \phi_2(nL) - 2\gamma F(\theta)\frac{\partial E\{F(\theta)\}}{\partial \phi_2}\bigg|_{\theta=\theta(nL)} =$$

15

$$= \phi_2(nL) - 2\gamma[y(t_{nL}) - u(nL)]\frac{\partial E\{y(t_{nL})\}}{\partial \phi_2} \qquad (8)$$

$$\theta_1[(n+1)L] = \theta_1(nL) - \gamma \frac{\partial E\{F^2(\theta)\}}{\partial \theta_1}\bigg|_{\theta=\theta(nL)} = \theta_1(nL) - 2\gamma F(\theta)\frac{\partial E\{F(\theta)\}}{\partial \theta_1}\bigg|_{\theta=\theta(nL)} =$$

$$= \theta_1(nL) - 2\gamma[y(t_{nL}) - u(nL)]\frac{\partial E\{y(t_{nL})\}}{\partial \theta_1}$$

20

$$\theta_2[(n+1)L] = \theta_2(nL) - \gamma \frac{\partial E\{F^2(\theta)\}}{\partial \theta_2}\bigg|_{\theta=\theta(nL)} = \theta_2(nL) - 2\gamma F(\theta)\frac{\partial E\{F(\theta)\}}{\partial \theta_2}\bigg|_{\theta=\theta(nL)} =$$

$$= \theta_2(nL) - 2\gamma[y(t_{nL}) - u(nL)]\frac{\partial E\{y(t_{nL})\}}{\partial \theta_2}$$

where $E\{.\}$ denotes "expected" and $\gamma > 0$ is a scale factor which controls the adjustment amount.

In vector notation this means that the vector of the compensator parameters is updated by adding a new vector with its norm proportional to the norm of the gradient $F^2(\theta)$ and with opposite direction, i.e. all its components have their sign changed:

$$\theta[(n+1)L] = \theta(nL) - \gamma \nabla E\{F^2(\theta)\}\big|_{\theta=\theta(nL)} = \theta(nL) - 2\gamma F(\theta)\nabla E\{F(\theta)\}\big|_{\theta=\theta(nL)} \quad (9)$$

This way we are sure to move towards a relative minimum of the function $F^2(\theta)$.

Three variations of the basic updating method defined in (8) can be obtained by using only sign information contained in the error $e(k)$ and/or the partial derivative. Hence the three possible variations are (considering for example the updating rule related to ϕ_1):

$$\varphi_1[(n+1)L] = \varphi_1(nL) \quad 2\gamma sign[y(t_{nL}) \quad u(nL)]\frac{\partial E\{y(t_{nL})\}}{\partial \varphi_1} \quad (10)$$

or

$$\varphi_1[(n+1)L] = \varphi_1(nL) \quad 2\gamma[y(t_{nL}) \quad u(nL)]sign \frac{\partial E\{y(t_{nL})\}}{\partial \varphi_1} \quad (11)$$

or

$$\varphi_1[(n+1)L] = \varphi_1(nL) \quad 2\gamma sign[y(t_{nL}) \quad u(nL)]sign \frac{\partial E\{y(t_{nL})\}}{\partial \varphi_1} \quad (12)$$

We shall now describe two methods of computing the gradient of the function $F^2(\theta)$ to find the required control parameters.

First method

Let us consider the updating rule in vector notation (9). To simplify this rule, in the error $F(\theta)=e(k)$ we substitute the transmitted information symbol $u(nL)$ with the corresponding decision $\hat{u}(nL)$, i.e. we substitute the error $e(k)$ with the estimated error $\hat{e}(k)$ defined as:

$$\hat{e}(k)=y(t_k)-\hat{u}(k) \tag{13}$$

In the diagram of FIG 1 this estimated error is obtained at the output of the subtractor block 20 and sent to the CD.

Defining $G[\theta(nL)]=E\{\hat{e}^2(nL)\}$, the updating rule (9) becomes:

$$\theta[(n+1)L] = \theta(nL) - \gamma\nabla G(\theta)\big|_{\theta=\theta(nL)} \tag{14}$$

The partial derivatives of $G(\theta)$ for $\theta=\theta(t_n)$ can be computed using the following 5-step procedure.

      –   Step 1: Find the value of $G[\theta(nL)]=G[\phi_1(nL),\ \phi_2(nL),\ \theta_1(nL),\ \theta_2(nL)]$ all'iterazione n. at iteration n. To do this, in the time interval $(nLT, nLT+LT/5)$ an estimate of $G[\theta(nL)]$ is computed by averaging the $L/5$ values of the estimated square error, i.e.,

$$G[\theta(nL)] = \frac{\sum_{i=0}^{L/5-1}\hat{e}(nL+i)}{L/5} \tag{15}$$

- Step 2: find the partial derivative:

$$\left.\frac{\partial G(\theta)}{\partial \phi_1}\right|_{\theta = \theta(nL)}$$

at iteration n. To do this, parameter ϕ_1 is set at $\phi_1(nL)+\Delta$ while the other parameters are left unchanged. The corresponding value of $G(\theta)$, i.e. $G[\phi_1(nL)+\Delta, \phi_2(nL), \theta_1(nL), \theta_2(nL)]$ is computed as in step 1 but in the time interval $(nLT+LT/5, nLT+2LT/5$. The estimate of the partial derivative of $G(\theta)$ as a function of ϕ_1 is computed as:

$$\left.\frac{\partial G(\theta)}{\partial \phi_1}\right|_{\theta = \theta(nL)} \cong \frac{G[\phi_1(nL) + \Delta, \phi_2(nL), \theta_1(nL), \theta_2(nL)] - G[\phi_1(nL), \phi_2(nL), \theta_1(nL), \theta_2(nL)]}{\Delta}$$

(16)

- Step 3. Find the partial derivative:

$$\left.\frac{\partial G(\theta)}{\partial \phi_2}\right|_{\theta = \theta(nL)}$$

at iteration n. To do this the parameter ϕ_2 is set at $\phi_2(nL)+\Delta$ while the other parameters are left changed. The corresponding value of $G(\theta)$, i.e. $G[\phi_1(nL), \phi_2(nL)+\Delta, \theta_1(nL), \theta_2(nL)]$, is computed as in step 1 but in the time interval $(nLT+2LT/5, nLT+3LT/5)$. The estimate of the partial derivative of $G(\theta)$ with respect to ϕ_2 is computed as:

$$\left.\frac{\partial G(\theta)}{\partial \phi_2}\right|_{\theta = \theta(nL)} \cong \frac{G[\phi_1(nL), \phi_2(nL) + \Delta, \theta_1(nL), \theta_2(nL)] - G[\phi_1(nL), \phi_2(nL), \theta_1(nL), \theta_2(nL)]}{\Delta}$$

(17)

12

-    Step 4: Find the partial derivative:

$$\frac{\partial G(\theta)}{\partial \theta_1}\bigg|_{\theta = \theta(nL)}$$

5   at iteration n. To do this, parameter θ_1 is set at $\theta_1(nL)+ \Delta$ while the other parameters are left unchanged. The corresponding value of $G(\theta)$, i.e. $G[\phi_1(nL)$, $\phi_2(nL)$, $\theta_1(nL)+\Delta$, $\theta_2(nL)]$, is computed as in Step 1 but in the time interval $(nLT+3LT/5, nLT+4LT/5)$. The estimate of the partial derivative of $G(\theta)$ with respect to $G(\theta)$ is computed as:

10

$$\frac{\partial G(\theta)}{\partial \theta_1}\bigg|_{\theta = \theta(nL)} \cong \frac{G[\phi_1(nL),\phi_2(nL),\theta_1(nL)+\Delta,\theta_2(nL)]- G[\phi_1(nL),\phi_2(nL),\theta_1(nL),\theta_2(nL)]}{\Delta}$$

(18)

-    Step 5: Find the partial derivative:

15

$$\frac{\partial G(\theta)}{\partial \theta_2}\bigg|_{\theta = \theta(nL)}$$

at iteration n. To do this the parameter θ_2 is set at $\theta_2(nL)+\Delta$ while the other parameters are left unchanged. The corresponding value of $G(\theta)$, i.e. $G[\phi_1(nL)$, $\phi_2(nL)$, $\theta_1(nL)$, $\theta_2(nL) +\Delta]$, is computed as in Step 1 but in the time interval

20    $(nLT+4LT/5, (n+1)LT)$. The estimate of the partial derivative of $G(\theta)$ with respect to θ_2 is computed as:

$$\frac{\partial G(\theta)}{\partial \theta_2}\bigg|_{\theta = \theta(nL)} \cong \frac{G[\phi_1(nL),\phi_2(nL),\theta_1(nL),\theta_2(nL)+\Delta]- G[\phi_1(nL),\phi_2(nL),\theta_1(nL),\theta_2(nL)]}{\Delta}$$

(19)

The above parameter update is done only after estimation of the gradient has been completed.

Note that in this case it is not necessary that the relationship between the control
5   parameters of PC and optical rotators and the corresponding Jones matrices be known.

Indeed, the partial derivatives of the function with respect to the compensator control parameters are computed without knowledge of this relationship. Consequently if the control parameters are different from those assumed as an example and are for example
10   some voltage or some other angle, we may similarly compute the partial derivative and update these different control parameters accordingly.

Lastly, it is noted that when this algorithm is used the CD is not necessary and the controller must receive the estimated error only.
15

Second Method

When an accurate characterization of the PC and each optical rotator is available the updating rules may be expressed as a function of the estimated error and the signals on
20   the two orthogonal polarizations at the compensator input.

Using the known stochastic gradient algorithm (for example as mentioned in the book Digital Communications by J.G.Proakis -McGraw-Hill, New York, 1983) and substituting in (8) the error e(nL) with the corresponding estimated error ê(nL) we find:

$$\phi_1[(n+1)L] = \phi_1(nL) - 2\gamma\hat{e}(nL)\frac{\partial E\{y(t_{nL})\}}{\partial \phi_1}$$

$$\phi_2[(n+1)L] = \phi_2(nL) - 2\gamma\hat{e}(nL)\frac{\partial E\{y(t_{nL})\}}{\partial \phi_2}$$

$$\theta_1[(n+1)L] = \theta_1(nL) - 2\gamma\hat{e}(nL)\frac{\partial E\{y(t_{nL})\}}{\partial \theta_1}$$ (20)

$$\theta_2[(n+1)L] = \theta_2(nL) - 2\gamma\hat{e}(nL)\frac{\partial E\{y(t_{nL})\}}{\partial \theta_2}$$

In vector notation the expression (9) becomes:

(21)

$$\theta[(n+1)L] = \theta(nL) - 2\gamma\hat{e}(nL)\nabla y(t_{nL})$$

Before describing how the gradient of $y(t_nL)$ is to be computed we introduce an equivalent model of the PMD compensator.

It was found that the PMD compensator shown in FIG. 1 can be modeled as an equivalent to a two-dimensional transverse filter using four tapped delay lines (TDL) combining the signals on the two principal polarization states (PSP). This equivalent model is shown in FIG 2 where:

$$c_1 \triangleq \cos\theta_1 \cos\theta_2 h_1$$
$$c_2 \triangleq -\sin\theta_1 \sin\theta_2 h_1$$
$$c_3 \triangleq -\sin\theta_1 \cos\theta_2 h_2^*$$
$$c_4 \triangleq -\cos\theta_1 \sin\theta_2 h_2^*$$
$$c_5 \triangleq \cos\theta_1 \cos\theta_2 h_2$$ (22)
$$c_6 \triangleq -\sin\theta_1 \sin\theta_2 h_2$$
$$c_7 \triangleq \sin\theta_1 \cos\theta_2 h_1^*$$
$$c_8 \triangleq \cos\theta_1 \sin\theta_2 h_1^*$$

For the sake of convenience let $c(\theta)$ designate the vector whose components are the c_i in (22). It is noted that the tap coefficients c_i of the four TDLs are not independent of each other. On the contrary, given four of them the others are completely determined by (22). In the FIG for the sake of clarity it is designated $\beta = 1 - \alpha$.

5

The partial derivatives of $y(t_n L)$ appearing in (21) may be expressed as a function of the components on the two PSPs of the signal at the compensator input at some appropriate instants. The output sample $y(t_k)$ may be written as (where $(B)^H$ indicates the transposed conjugate of the matrix B):

10

$$y(t_k) = c^H A(k) c \qquad (23)$$

where the Hermitian matrix $A(k)$ is given by:

15

$$A(k) = a(k)^* a^T(k) + b(k)^* b^T(k) \qquad (24)$$

with vectors $a(k)$ and $b(k)$ defined by:

20

$$a(k) = \begin{pmatrix} x_1(t_k) \\ x_1(t_k - \alpha\tau_c) \\ x_1(t_k - \tau_c) \\ x_1(t_k - \tau_c - \alpha\tau_c) \\ x_2(t_k) \\ x_2(t_k - \alpha\tau_c) \\ x_2(t_k - \tau_c) \\ x_2(t_k - \tau_c - \alpha\tau_c) \end{pmatrix} \qquad b(t_k) = \begin{pmatrix} x_2^*(t_k - 2\tau_c) \\ x_2^*(t_k - \tau_c - \beta\tau_c) \\ x_2^*(t_k - \tau_c) \\ x_2^*(t_k - \beta\tau_c) \\ -x_1^*(t_k - 2\tau_c) \\ -x_1^*(t_k - \tau_c - \beta\tau_c) \\ -x_1^*(t_k - \tau_c) \\ -x_1^*(t_k - \beta\tau_c) \end{pmatrix} \qquad (25)$$

Computing the gradient of $y(t_nL)$, the algorithm (9) becomes:

$$\theta[(n+1)L]=\theta(nL)-4\gamma\,\hat{e}(nL)\mathrm{Re}\{J^HA(nL)c\} \tag{26}$$

5   where

$$J \triangleq \begin{pmatrix} \dfrac{\partial c_1}{\partial \phi_1} & \dfrac{\partial c_1}{\partial \phi_2} & \dfrac{\partial c_1}{\partial \theta_1} & \dfrac{\partial c_1}{\partial \theta_2} \\ \dfrac{\partial c_2}{\partial \phi_1} & \dfrac{\partial c_2}{\partial \phi_2} & \dfrac{\partial c_2}{\partial \theta_1} & \dfrac{\partial c_2}{\partial \theta_2} \\ \vdots & \vdots & \vdots & \vdots \\ \dfrac{\partial c_8}{\partial \phi_1} & \dfrac{\partial c_8}{\partial \phi_2} & \dfrac{\partial c_8}{\partial \theta_1} & \dfrac{\partial c_8}{\partial \theta_2} \end{pmatrix} \tag{27}$$

10

is the Jacobean matrix of the transformation $c=c(\theta)$.

When the control parameters are different from those taken as examples we will have different relationships between these control parameters and coefficients c_i. For

15   example, if the PC is controlled by means of some voltages, given the relationship between these voltages and the coefficients h_1 and h_2 which appear in (3) by using the equations (22) we will always be able to express the coefficients ci as a function of these new control parameters.

20   Consequently in computing the gradient of $y(t_nL)$ the only change we have to allow for is the expression of the Jacobean matrix J, which has to be changed accordingly as readily imaginable to those skilled in the art.

Lastly it is noted that when this second method is used the CD must receive the optical signals at the input of the compensator and the estimated error $\hat{e}(nL)$. The CD must supply the controller directly with this estimated error and with the signal samples $x_1(t)$ and $x_2(t)$ at the desired instants.

5

It is now clear that the predetermined purposes have been achieved by making available an effective method for adaptive control of a PMD compensator and a compensator applying this method.

10 Naturally the above description of an embodiment applying the innovative principles of the present invention is given by way of non-limiting example of said principles within the scope of the exclusive right claimed here.

CLAIMS

1.   Method for the adaptive adjustment of a PMD compensator in optical fiber communication systems with the compensator comprising a cascade of adjustable optical devices over which passes an optical signal to be compensated comprising the steps of:

   a.   extracting the $y_1(t)$ and $y_2(t)$ on the two orthogonal polarizations of the signal at the compensator output,

   b.   obtaining the signal $y(t)=|y_1(t)|^2+|y_2(t)|^2$;

   c.   sampling the signal $y(t)$ at instants $t_k=kT$ with T = symbol interval to obtain samples $y(t_k)$,

   d.   computing the mean square error $e(k)=y(tk)-u(k)$ with $u(k)$ equal to the symbol transmitted or with $u(k)$ replaced by a decision $\hat{u}(k)$ on the transmitted symbol $u(k)$, and

   e.   producing control signals for parameters of at least some of said adjustable optical devices to tend toward minimization of $e(k)$.

2.   Method in accordance with claim 1 in which said parameters being consolidated in a vector θ and a function $F(\theta)=e(k)$ being defined the parameters are adjusted to tend to minimize the mean value of $F^2(\theta)$.

3.   Method in accordance with claim 2 in which the vector θ being defined as:
$$\theta \triangleq (\phi_1,\phi_2,\theta_1,\theta_2)^T$$

where ϕ_1, ϕ_2, θ_1 and θ_2 are said parameters these parameters are updated by the rule:

$$\phi_1[(n+1)L] = \phi_1(nL) - \gamma \frac{\partial E\{F^2(\theta)\}}{\partial \phi_1}\bigg|_{\theta = \theta(nL)} = \phi_1(nL) - 2\gamma F(\theta)\frac{\partial E\{F(\theta)\}}{\partial \phi_1}\bigg|_{\theta = \theta(nL)} =$$

$$= \phi_1(nL) - 2\gamma [y(t_{nL}) - u(nL)]\frac{\partial E\{y(t_{nL})\}}{\partial \phi_1}$$

$$\phi_2[(n+1)L] = \phi_2(nL) - \gamma \frac{\partial E\{F^2(\theta)\}}{\partial \phi 2}\bigg|_{\theta = \theta(nL)} = \phi_2(nL) - 2\gamma F(\theta)\frac{\partial E\{F(\theta)\}}{\partial \phi_2}\bigg|_{\theta = \theta(nL)} = \qquad (8)$$

$$= \phi_2(nL) - 2\gamma [y(t_{nL}) - u(nL)]\frac{\partial E\{y(t_{nL})\}}{\partial \phi_2}$$

$$\theta_1[(n+1)L] = \theta_1(nL) - \gamma \frac{\partial E\{F^2(\theta)\}}{\partial \theta_1}\bigg|_{\theta = \theta(nL)} = \theta_1(nL) - 2\gamma F(\theta)\frac{\partial E\{F(\theta)\}}{\partial \theta_1}\bigg|_{\theta = \theta(nL)} =$$

$$= \theta_1(nL) - 2\gamma [y(t_{nL}) - u(nL)]\frac{\partial E\{y(t_{nL})\}}{\partial \theta_1}$$

$$\theta_2[(n+1)L] = \theta_2(nL) - \gamma \frac{\partial E\{F^2(\theta)\}}{\partial \theta_2}\bigg|_{\theta = \theta(nL)} = \theta_2(nL) - 2\gamma F(\theta)\frac{\partial E\{F(\theta)\}}{\partial \theta_2}\bigg|_{\theta = \theta(nL)} =$$

$$= \theta_2(nL) - 2\gamma [y(t_{nL}) - u(nL)]\frac{\partial E\{y(t_{nL})\}}{\partial \theta_2}$$

where E{.} denotes "expected" and γ>0 is a scale factor which controls the adjustment amount.

4.   Method in accordance with claim 2 in which the vector θ of the parameters is updated by adding a new vector with the norm proportionate to the norm of the gradient of $F^2(\theta)$ and with opposite direction, i.e. all its components have their sign changed so that the updating rule is:

$$\theta[(n+1)L] = \theta(nL) - \gamma \nabla E\{F^2(\theta)\}\big|_{\theta = \theta(nL)} = \theta(nL) - 2\gamma F(\theta)\nabla E\{F(\theta)\}\big|_{\theta = \theta(nL)} \qquad (9)$$

so that movement is towards a relative minimum of the function $F^2(\theta)$.

5.   Method in accordance with claim 2 in which in (8) only the sign of the error e(k)=y(t_k)-u(k) and/or of the partial derivative is used.

6.   Method in accordance with claim 4 in which in the error $F(\theta)=e(k)$ the transmitted information symbol u(nL) is substituted with the corresponding decision û(nL) so as to substitute the error e(k) with the estimated error ê(k) defined as ê(k)=y(t_k)-û(k).

7.   Method in accordance with claim 6 in which $G[\theta(nL)]=E\{ê^2(nL)\}$ is defined so that the updating rule (9) becomes:

$$\theta[(n+1)L] = \theta(nL) - \gamma \nabla G(\theta)\big|_{\theta = \theta(nL)} \tag{14}$$

8.   Method in accordance with claim 7 in which the partial derivatives of $G(\theta)$ for $\theta=\theta(t_n)$ are computed by the following five-step procedure:

—   **Step 1.** Find the value of $G[\theta(nL)]=G[\phi1(nL), \phi2(nL), \theta1(nL), \theta2(nL)]$ at iteration n; to do this, in the time interval (nLT, nLT+LT/5) an estimate of $G[\theta(nL)]$ is computed by averaging the L/5 values of the estimated square error, i.e.,

$$G[\theta(nL)] = \frac{\sum_{i=0}^{L/5-1} ê(nL+i)}{L/5} \tag{15}$$

—   **Step 2.** Find the partial derivative:

$$\frac{\partial G(\theta)}{\partial \phi_1}\bigg|_{\theta = \theta(nL)}$$

at iteration n; to do this, parameter ϕ_1 is set at $\phi_1(nL)+\Delta$ while the other

parameters are left unchanged; the corresponding value of $G(\theta)$, i.e.

$G[\phi_1(nL)+\Delta, \phi_2(nL), \theta_1(nL), \theta_2(nL)]$ is computed as in step 1 but in the time

interval $(nLT+LT/5, nLT+2LT/5)$; the estimate of the partial derivative of

$G(\theta)$ with respect to ϕ_1 is computed as:

$$\left.\frac{\partial G(\theta)}{\partial \phi_1}\right|_{\theta = \theta(nL)} \cong \frac{G[\phi_1(nL) + \Delta, \phi_2(nL), \theta_1(nL), \theta_2(nL)] - G[\phi_1(nL), \phi_2(nL), \theta_1(nL), \theta_2(nL)]}{\Delta}$$

(16)

- **Step 3.** Find the partial derivative:

$$\left.\frac{\partial G(\theta)}{\partial \phi_2}\right|_{\theta = \theta(nL)}$$

at iteration n; to do this the parameter ϕ_2 is set at $\phi_2(nL)+\Delta$ while the other

parameters are left unchanged; the corresponding value of $G(\theta)$, i.e.

$G[\phi_1(nL), \phi_2(nL)+\Delta, \theta_1(nL), \theta_2(nL)],)]$, is computed as in step 1 but in the

time interval $(nLT+2LT/5, nLT+3LT/5)$; the estimate of the partial

derivative of $G(\theta)$ with respect to ϕ_2 is computed as:

$$\left.\frac{\partial G(\theta)}{\partial \phi_2}\right|_{\theta = \theta(nL)} \cong \frac{G[\phi_1(nL), \phi_2(nL) + \Delta, \theta_1(nL), \theta_2(nL)] - G[\phi_1(nL), \phi_2(nL), \theta_1(nL), \theta_2(nL)]}{\Delta}$$

(17)

-   **Step 4.** Find the partial derivative:

$$\frac{\partial G(\theta)}{\partial \theta_1}\bigg|_{\theta = \theta(nL)}$$

at iteration n; to do this, parameter θ_1 is set at $\theta_1(nL) + \Delta$ while the other parameters are left unchanged; the corresponding value of $G(\theta)$, i.e. $G[\phi_1(nL), \phi_2(nL), \theta_1(nL) + \Delta, \theta_2(nL)],)]$, is computed as in Step 1 but in the time interval $(nLT + 3LT/5, nLT + 4LT/5)$; the estimate of the partial derivative of $G(\theta)$ with respect to $G(\theta)$ is computed as:

$$\frac{\partial G(\theta)}{\partial \theta_1}\bigg|_{\theta = \theta(nL)} \cong \frac{G[\phi_1(nL), \phi_2(nL), \theta_1(nL) + \Delta, \theta_2(nL)] - G[\phi_1(nL), \phi_2(nL), \theta_1(nL), \theta_2(nL)]}{\Delta}$$

(13)

-   **Step 5.** Find the partial derivative:

$$\frac{\partial G(\theta)}{\partial \theta_2}\bigg|_{\theta = \theta(nL)}$$

at iteration n; to do this the parameter θ_2 is set at $\theta_2(nL) + \Delta$ while the other parameters are left unchanged; the corresponding value of $G(\theta)$, i.e. $G[\phi_1(nL), \phi_2(nL), \theta_1(nL), \theta_2(nL) + \Delta]$, is computed as in Step 1 but in the time interval $(nLT + 4LT/5, (n+1)LT)$; the estimate of the partial derivative of $G(\theta)$ with respect to θ_2 is computed as:

$$\frac{\partial G(\theta)}{\partial \theta_2}\bigg|_{\theta = \theta(nL)} \cong \frac{G[\phi_1(nL), \phi_2(nL), \theta_1(nL), \theta_2(nL) + \Delta] - G[\phi_1(nL), \phi_2(nL), \theta_1(nL), \theta_2(nL)]}{\Delta}$$

(14)

9. Method in accordance with claim 3 in which in (8) the error e(nL) is substituted with the estimated error ê(nL) and the updating rules become:

$$\phi_1[(n+1)L] = \phi_1(nL) - 2\gamma\hat{e}(nL)\frac{\partial E\{y(t_{nL})\}}{\partial \phi_1}$$

$$\phi_2[(n+1)L] = \phi_2(nL) - 2\gamma\hat{e}(nL)\frac{\partial E\{y(t_{nL})\}}{\partial \phi_2}$$

$$\theta_1[(n+1)L] = \theta_1(nL) - 2\gamma\hat{e}(nL)\frac{\partial E\{y(t_{nL})\}}{\partial \theta_1} \qquad (20)$$

$$\theta_2[(n+1)L] = \theta_2(nL) - 2\gamma\hat{e}(nL)\frac{\partial E\{y(t_{nL})\}}{\partial \theta_2}$$

10. Method in accordance with claim 4 in which the error e(nL) is substituted with the estimated error ê(nL) so that the updating rule (9) becomes:

$$\theta[(n+1)L] = \theta(nL) - 2\gamma\hat{e}(nL)\nabla y(t_{nL}) \qquad (21)$$

11. Method in accordance with claim 1 in which the PMD compensator is modeled as a two-dimensional transverse filter using four tapped delay lines (TDL) combining the signals on the two principal polarization states (PSP).

12. Method in accordance with claims 2 and 11 in which the vector θ of said parameters is updated by adding a new vector with the norm proportionate to the norm of the gradient of $F^2(\theta)$ and with opposite direction, i.e. all its components have their sign changed so that the updating rule is:

$$\theta[(n+1)L] = \theta(nL) - 4\gamma \hat{e}(nL)\text{Re}\{J^H A(nL)c\} \qquad (26)$$

with the Hermitian matrix A(k) given by:

$$A(k)=a(k)^{*}a^{T}(k)+b(k)^{*}b^{T}(k) \tag{24}$$

with the vectors a(k) e b(k) given by:

$$a(k) = \begin{pmatrix} x_1(t_k) \\ x_1(t_k - \alpha\tau_c) \\ x_1(t_k - \tau_c) \\ x_1(t_k - \tau_c - \alpha\tau_c) \\ x_2(t_k) \\ x_2(t_k - \alpha\tau_c) \\ x_2(t_k - \tau_c) \\ x_2(t_k - \tau_c - \alpha\tau_c) \end{pmatrix} \qquad b(t_k) = \begin{pmatrix} x_2^*(t_k - 2\tau_c) \\ x_2^*(t_k - \tau_c - \beta\tau_c) \\ x_2^*(t_k - \tau_c) \\ x_2^*(t_k - \beta\tau_c) \\ -x_1^*(t_k - 2\tau_c) \\ -x_1^*(t_k - \tau_c - \beta\tau_c) \\ -x_1^*(t_k - \tau_c) \\ -x_1^*(t_k - \beta\tau_c) \end{pmatrix} \tag{25}$$

and

$$\mathbf{J} \triangleq \begin{pmatrix} \dfrac{\partial c_1}{\partial \phi_1} & \dfrac{\partial c_1}{\partial \phi_2} & \dfrac{\partial c_1}{\partial \theta_1} & \dfrac{\partial c_1}{\partial \theta_2} \\ \dfrac{\partial c_2}{\partial \phi_1} & \dfrac{\partial c_2}{\partial \phi_2} & \dfrac{\partial c_2}{\partial \theta_1} & \dfrac{\partial c_2}{\partial \theta_2} \\ \vdots & \vdots & \vdots & \vdots \\ \dfrac{\partial c_8}{\partial \phi_1} & \dfrac{\partial c_8}{\partial \phi_2} & \dfrac{\partial c_8}{\partial \theta_1} & \dfrac{\partial c_8}{\partial \theta_2} \end{pmatrix} \tag{27}$$

with $c_1,....,c_8$ which are the tap coefficients of the four tapped delay lines and

x1(t), x2(t) are the components of the two principal polarization states at the

compensator input.

13.   Method in accordance with claim 1 in which said optical devices comprise a polarization controller with control angles ϕ_1, ϕ_2 and two optical rotators with rotation angles θ_1 and θ_2 and said parameters comprise said control angles ϕ_1, ϕ_2 and said rotation angles θ_1, θ_2 or functions thereof.

14.   Method in accordance with claim 13 in which between the controller and an optical rotator and between optical rotators there are fibers which introduce a predetermined differential unit delay maintaining the polarization.

15.   Compensator for PMD in optical fiber communication systems applying the method in accordance with any one of the above claims and comprising a cascade of adjustable optical devices over which passes an optical signal to be compensated and an adjustment system comprising a photodetector (17) which takes the components $y_1(t)$ and $y_2(t)$ on the two orthogonal polarizations from the signal at the compensator output, a sampler (19) which samples at instants $t_k=kT$ with T = symbol interval, the signal $y(t)=|y_1(t)|^2+|y_2(t)|^2$ at the output of the photodetector (17) to obtain samples $y(t_k)$, a circuit (18,20) for computation of the mean square error $e(k)=y(t_k)-u(k)$ with $u(k)$ equal to the symbol transmitted and a regulator (15,16) which regulates parameters of at least some of said optical devices to tend towards minimization of $e(k)$.

16.   Compensator in accordance with claim 15 characterized in that said optical devices comprise a polarization controller with control angles ϕ_1, ϕ_2 and two optical rotators with rotation angles θ_1 and θ_2 and in which said parameters which are adjusted consist of said control angles ϕ_1, ϕ_2 and said rotation angles θ_1 and θ_2.

17.   Compensator in accordance with claim 16 characterized in that between the controller and optical rotator and between optical rotators there are fibers which introduce a predetermined differential unit delay maintaining the polarization.
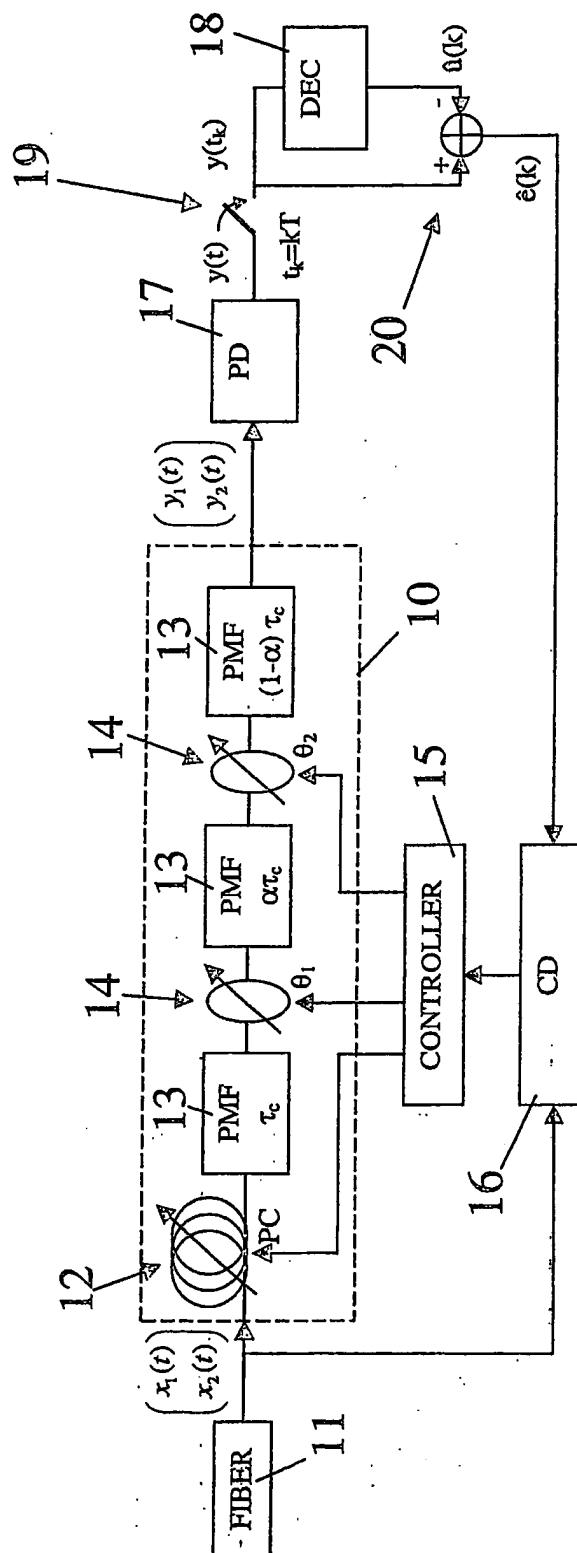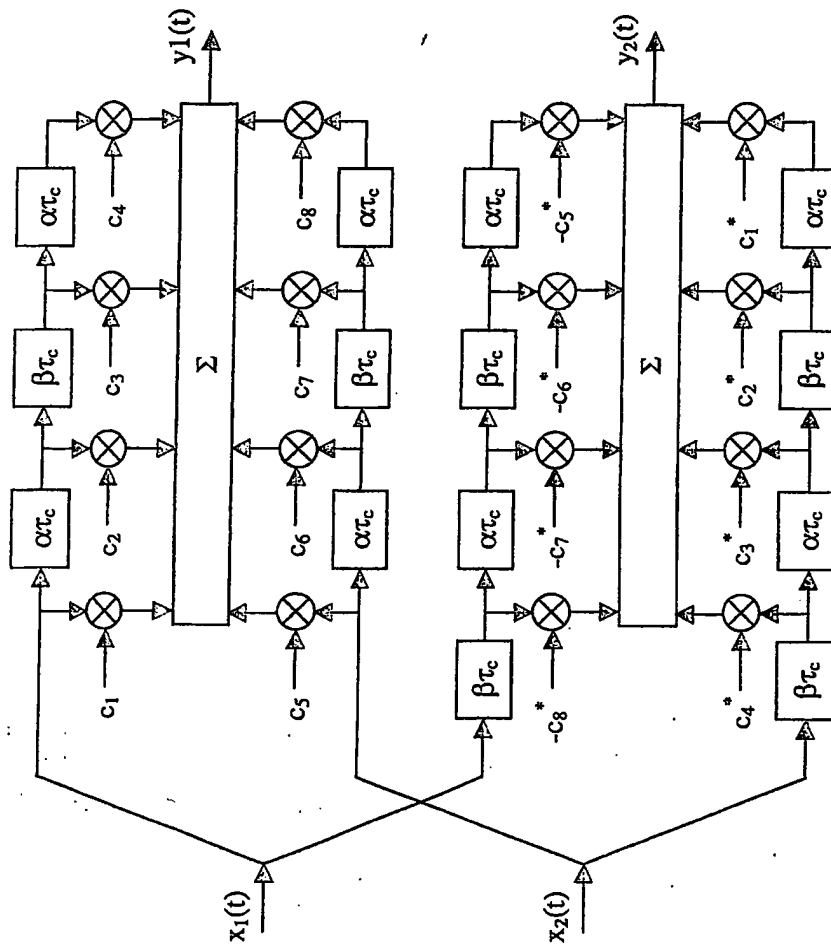
Fig. 1

Fig.2

INTERNATIONAL SEARCH REPORT

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7   H04B10/18   G02B6/34

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7   H04B   G02B

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data, PAJ

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	BUELOW H ET AL: "ADAPTIVE PMD MITIGATION AT 10 GBIT/S USING AN ELECTRONIC SIGE EQUALISER IC" 25TH EUROPEAN CONFERENCE ON OPTICAL COMMUNICATION. (ECOC'99). NICE, FRANCE, SEPT. 27 - 30, 1999. REGULAR AND INVITED PAPERS, EUROPEAN CONFERENCE ON OPTICAL COMMUNICATION (ECOC), PARIS: SEE, FR, vol. II OF II, 26 September 1999 (1999-09-26), pages II-138-II-139, XP001035418 ISBN: 2-912328-12-8 the whole document --- -/--	1-4,6, 11,15

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

'A' document defining the general state of the art which is not considered to be of particular relevance

'E' earlier document but published on or after the international filing date

'L' document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

'O' document referring to an oral disclosure, use, exhibition or other means

'P' document published prior to the international filing date but later than the priority date claimed

'T' later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

'X' document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

'Y' document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

'&' document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
28 March 2003	03/04/2003

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31–70) 340–2040, Tx. 31 651 epo nl, Fax: (+31–70) 340–3016	Reville, L

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
Y A	US 6 115 419 A (MEEHAN JOSEPH P) 5 September 2000 (2000-09-05) figures 1,4 column 7, line 10 –column 9, line 22 ---	1-4,6, 11,15 7-10,12
A	SANDEL D ET AL: "Automatic polarisation mode dispersion compensation in 40 Gbit/s optical transmission system" ELECTRONICS LETTERS, IEE STEVENAGE, GB, vol. 34, no. 23, 12 November 1998 (1998-11-12), pages 2258-2259, XP006010567 ISSN: 0013-5194 the whole document ---	13,14, 16,17
A	EP 0 725 477 A (NIPPON TELEGRAPH & TELEPHONE) 7 August 1996 (1996-08-07) abstract column 10, line 52 –column 11, line 19 figures 6,7 ———	11

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

Patent document cited in search report		Publication date	Patent family member(s)		Publication date
US 6115419	A	05-09-2000	WO	0129987 A1	26-04-2001
			EP	1142160 A1	10-10-2001
EP 0725477	A	07-08-1996	CN	1136751 A ,B	27-11-1996
			DE	69614284 D1	13-09-2001
			DE	69614284 T2	10-01-2002
			EP	0725477 A2	07-08-1996
			JP	3239926 B2	17-12-2001
			JP	8274664 A	18-10-1996
			US	5675613 A	07-10-1997